PRODUCTION TITLE PAGE

Feature	**eResources**
Title	**Ontology in Immunology**
Authors	Catharien Hilkens, PhD Phillip Lord, PhD
Corr address – please use this address for correspondence	catharien.hilkens@newcastle.ac.uk phillip.lord@newcastle.ac.uk
Disclosure:	The authors declare no conflicts of interest.
Special feature: Link library	
Figure is a stock image	**Every month use the same figure. It should be centered between the columns as laid out on template.**

Article

Overcoming the barrier of graft rejection by the immune system has always been a major concern for the transplantologist. Since the seminal work of Medawar & colleagues on acquired immunological tolerance to allogeneic skin transplants in the 1950s (1), the discovery of functionally distinct immune cell subsets has shaped the field of Transplant Immunology. We now have a much better picture of how immune cells are involved in graft rejection and how we may prevent this. Paradoxically, the more knowledge we gain about the functional properties of immune cells, the more we are struggling to define these cells, which hampers further progress in the field.

While we all think we know what the definition of a T-cell is, different researchers will consider different aspects to be of importance: they might mention the thymus, cell-mediated immunity, the T-cell receptor. If you do not believe us, try asking the next five immunologists you meet for their definition. As we move to subtypes, these differences will turn to disagreements. What evidence should we accept to demonstrate that a cell is a naturally occurring or an adaptive regulatory T-cell, or, dare we say, a dendritic cell or a macrophage? If we do not agree on this, it is hard to make our science reproducible.

Good definitions are surprisingly hard to make and, in fact, there is an entire discipline devoted to it called ‘ontology’. This has become more important in many areas as they move toward big data approaches – computers require accurate definitions if they are to

support science (2). Here, we interpret "ontology" broadly to mean, any attempt to organize and standardize the way we describe and represent our knowledge, so enabling different scientists to accurately compare their results with each other.

In this article, we will describe several online resources that are useful for the readership of Transplantation in organizing or describing their data, and expanding this with new knowledge. These resources include minimum information models (MIMs), databases and ontologies.

MIMs are reporting guidelines that can be used to ensure that all the critical information is available in published data. This increases the transparency and reusability of these data. As such, MIMs are an important first step toward standardization of experimental procedures. The first MIM in the biomedical field was MIAME - Minimum Information About a Microarray Experiment - [**A**] (3). In the field of immunology, Minimal Information About T cell Assays (MIATA)[**B**] provides a framework for the reporting of experiments that measure the characteristics and function of T cells (4). MIATA was originally designed for the immunomonitoring field, aiming to increase the quality and comparability of T cell biomarker data obtained in clinical trials and studies. It is also a useful resource for the reporting of T cell data from non-clinical research studies. Another recent initiative is in the field of tolerance-inducing antigen-presenting cell (tolAPC) therapies, which have been developed for the prevention of graft rejection after transplantation, or for treatment of autoimmune diseases (5). The Minimum Information about Tolerogenic Antigen-Presenting cells (MITAP) reporting guidelines [**C**] are very timely considering a number of tolAPC products are currently being tested or have been tested in clinical trials (6). It will facilitate a better comparison of the differences and similarities of these therapeutic cells.

A clear example of a standardized nomenclature that facilitates better communication between scientists is the numbering of surface molecules on immune cells. The Human Leucocyte Differentiation Antigen (HLDA) workshops have taken place 10 times since 1980, identifying new cell surface molecules and assigning them a cluster of differentiation (CD) number (7). The full list and new entries can be found on the Human Cell Differentiation Molecules (HCDM) website [**D**].

The Gene Ontology (GO - [E]) is an ontology in the more strict sense of the word: classifications that enable searching and description. GO classifies the function of gene products at the level of i) molecular activity; ii) biological processes and pathways and iii) the cellular components in which they are active (8). Importantly, GO describes gene product characteristics in a species-independent manner, facilitating comparisons cross-species. By providing a controlled vocabulary for gene product attributes, published data can be annotated in a standardized and computer-readable way.

Probably more directly relevant is the specialist database, IPD-IMGT/HLA, which includes sequence for the human MHC (9) [F]. It is part of the international ImMunoGeneTics (IMGT) information system, which is a resource describing immunoglobulin and T-cell receptor genes (10) [G]. It includes nucleotide and protein sequences, polymorphisms and a database of therapeutic antibodies and fusion proteins. In addition to its suite of databases, it also includes many tools accessible through its website.

The Immune Epitope DataBase (IEDB) [H] is a significant resource containing data extracted manually from 15,000 journal articles published since 1960 (11). Their website allows searching for epitopes by name, species or disease. The same team also provides a set of tools for epitope prediction from a peptide sequence.

Both the IMGT and IEDB are driven underneath by a number of ontologies. The IMGT ontology, for example, provides a standard nomenclature for IG, TR and MHC proteins. Likewise, the IEDB uses the MHC restriction ontology [I]. The key problem that this addresses is the heterogeneity of naming between different species. The MHC restriction ontology builds on resources including the IMGT ontologies, and attempts to harmonize and cross-link terminology cross-species (12).

The final problem to be answered is how to find new e-resources. One solution here is to read a review article, but an alternative can be found at biosharing.org [J], which is a comprehensive resource describing the (many!) MIMs, databases and ontologies that exist.

Defining and naming things has always been difficult and has caused controversy in many different areas of biology – consider the classic paper "What, if anything, is a rabbit?"(13). As we get more complex, with more cell types and more protocols, this is becoming a pressing issue for immunology also; the reproducibility crisis may not have hit us yet, but (like Brexit) it is looming. Currently ontologies are more often a resource for database builders and informaticians than immunology researchers. However, if we wish to share, compare and reuse data between different laboratories, we will inevitably need to organise and catalogue our data better. They will not solve all the issues of reproducibility, but they will at least provide us with a common language, and that is a start.

Links

[A] http://fged.org/projects/miame/

[B] http://miataproject.org/

[C] http://w3id.org/ontolink/mitap

[D] http://www.hcdm.org/

[E] http://geneontology.org

[F] http://www.ebi.ac.uk/ipd/imgt/hla/

[G] http://www.imgt.org

[H] http://www.iedb.org

[I] http://purl.bioontology.org/ontology/MHC

[J] https://biosharing.org

Acknowledgements

We would like to thank Bjoern Peters for his feedback on this manuscript.

References

1. Billingham RE, Brent L, Medawar PB. Actively acquired tolerance of foreign cells. *Nature.* 1953;172:603-6.

2. R. Stevens and P. Lord. Application of ontologies in bioinformatics. In S. Staab and R. Studer, editors, *Handbook on Ontologies in Information Systems*. Springer, second edition, 2008.

3. Brazma A, Hingamp P, Quackenbush J, et al. Minimum information about a microarray

experiment (MIAME)-toward standards for microarray data. *Nat Genet.* 2001;29;365-71.

4. Britten CM, Janetzki S, van der Burg SH, et al. Minimal information about T cell assays: the process of reaching the community of T cell immunologists in cancer and beyond. *Cancer*
Immunol Immunother. 2011;60:15-22.

5. Ten Brinke A, Hilkens CM, Cools N, et al. Clinical Use of Tolerogenic Dendritic Cells-Harmonization Approach in European Collaborative Effort. *Mediators Inflamm.* 2015;2015:471719. doi: 10.1155/2015/471719.

6. Lord P, Spiering R, Aguillon JC, et al. Minimum information about tolerogenic antigen-presenting cells (MITAP): a first step towards reproducibility and standardisation of cellular therapies. *PeerJ*. 2016;in press.

7. Engel P, Boumsell L, Balderas R, et al. CD Nomenclature 2015: Human Leukocyte Differentiation Antigen Workshops as a Driving Force in Immunology*. J Immunol.* 2015;195:4555-63.

8. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000;25:25-9.

9. Robinson J, Halliwell JA, Hayhurst JH, et al. The IPD and IPD-IMGT/HLA Database: allele variant databases *Nucleic Acids Res.* 2015;43:D423-431

10. Lefranc MP, Giudicelli V, Duroux P, et al. IMGT®, the international ImMunoGeneTics information system® 25 years on. *Nucleic Acids Res.* 2015:43(Database issue):D413-22.

11. Vita R, Overton JA, Greenbaum JA, et al. The immune epitope database (IEDB) 3.0. *Nucleic Acids Res.* 2015:43(Database issue):D405-12.

12. Vita R, Overton JA, Seymour E, et al. An ontology for major histocompatibility restriction. *J Biomed Semantics.* 2016;11:7:1.

13. Wood, AE. What, if anything, is a rabbit? *Evolution.* 1957;4:417-25.